05-90
1205

2

```
RAW SEQUENCE LISTING              DATE: 12/26/2001
PATENT APPLICATION:  US/09/896,032       TIME: 15:29:49

Input Set : N:\Crf3\RULE60\09896032.txt
Output Set: N:\CRF3\12262001\I896032.raw

        SEQUENCE LISTING
 1 (1) GENERAL INFORMATION:
 2      (i) APPLICANT: Seidel, Christoph; Weinhues, Ursula-Henrike;
 3                     Schmitt, Urban; Motz, Manfred; Wiedmann, Michael;
 4                     Upmeier, Barbara; Soutscheck, Erwin
 5     (ii) TITLE OF INVENTION: Recombinant antigen from the NS3 region of
 6                     the hepatitis C virus
 7    (iii) NUMBER OF SEQUENCES: 9
 8     (iv) CORRESPONDENCE ADDRESS:
 9          (A) ADDRESSEE: Felfe & Lynch
10          (B) STREET: 805 Third Avenue
11          (C) CITY: New York
12          (D) STATE: New York
13          (E) COUNTRY: USA
14          (F) ZIP: 10022
15      (v) COMPUTER READABLE FORM:
16          (A) MEDIUM TYPE: 3.5 inch 1.44 Mb storage diskette
17          (B) COMPUTER: IBM PS/2
18          (C) OPERATING SYSTEM: PC-DOS
19          (D) SOFTWARE: PatentIn Relase #1.0,
20                     Version #1.25 (EPA)
21     (vi) CURRENT APPLICATION DATA:
22          (A) APPLICATION NUMBER: US/09/896,032
23          (B) FILING DATE: 29-Jun-2001
24          (C) CLASSIFICATION:
26    (vii) PRIOR APPLICATION DATA:
27          (A) APPLICATION NUMBER: 08/892,704
28          (B) FILING DATE:
31   (viii) ATTORNEY/AGENT INFORMATION:
32          (A) NAME: HANSON, NORMAN
33          (B) REGISTRATION NUMBER: 30, 946
34          (C) REFERENCE/DOCKET NUMBER: HUBR 1067
35     (ix) TELECOMMUNICATION INFORMATION:
36          (A) TELEPHONE: 212-688-9200
37          (B) TELEFAX: 212-838-3884
38 (2) INFORMATION FOR SEQ ID NO: 1:
39      (i) SEQUENCE CHARACTERISTICS:
40          (A) LENGTH: 885 base pairs
41          (B) TYPE: nucleic acid
42          (C) STRANDEDNESS: both
43          (D) TOPOLOGY: linear
44     (ii) MOLECULE TYPE: cDNA
45     (vi) ORIGINAL SOURCE:
46          (A) ORGANISM: hepatitis C virus
47   (viii) POSITION IN GENOME:
48          (A) CHROMOSOME/SEGMENT: NS3
49      (ix) FEATURE:
```

C--> 22
C--> 23

ENTERED

Input Set : N:\Crf3\RULE60\09896032.txt
Output Set: N:\CRF3\12262001\I896032.raw

```
50          (A) NAME/KEY: CDS
51          (B) LOCATION: 1..885
52     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:
53     ATG ACC ATG ATT ACG AAT TCC CGG GGA TCC ATC ATG AAA TCC CCG GTG        48
54     Met Thr Met Ile Thr Asn Ser Arg Gly Ser Ile Met Lys Ser Pro Val
55      1               5                  10                  15
56     TTC ACG GAT AAC TCC TCT CCA CCG GTA GTG CCC CAG AGC TTC CAG GTG        96
57     Phe Thr Asp Asn Ser Ser Pro Pro Val Val Pro Gln Ser Phe Gln Val
58              20                  25                  30
59     GCT CAC CTG CAT GCT CCC ACA GGC AGC GGC AAG AGC ACC AAG GTC CCG       144
60     Ala His Leu His Ala Pro Thr Gly Ser Gly Lys Ser Thr Lys Val Pro
61              35                  40                  45
62     GCT GCA TAC GCA GCT CAG GGC TAC AAG GTG CTA GTG CTC AAC CCT TCT       192
63     Ala Ala Tyr Ala Ala Gln Gly Tyr Lys Val Leu Val Leu Asn Pro Ser
64          50                  55                  60
65     GTT GCT GCA ACA TTG GGC TTT GGT GCC TAC ATG TCC AAG GCT CAT GGG       240
66     Val Ala Ala Thr Leu Gly Phe Gly Ala Tyr Met Ser Lys Ala His Gly
67      65                  70                  75                  80
68     ATC GAT CCT AAC ATC AGG ACC GGG GTG AGA ACA ATT ACC ACT GGC AGC       288
69     Ile Asp Pro Asn Ile Arg Thr Gly Val Arg Thr Ile Thr Thr Gly Ser
70                  85                  90                  95
71     CCC ATT ACG TAC TCC ACT TAC GGC AAG TTT CTT GCC GAC GGC GGG TGC       336
72     Pro Ile Thr Tyr Ser Thr Tyr Gly Lys Phe Leu Ala Asp Gly Gly Cys
73              100                 105                 110
74     GCA GGG GGT GCT TAT GAC ATA ATA ATT TGT GAC GAG TGC CAC TCC ACG       384
75     Ala Gly Gly Ala Tyr Asp Ile Ile Ile Cys Asp Glu Cys His Ser Thr
76              115                 120                 125
77     GAT GCC ACA TCC ATC TTG GGC ATC GGC ACT GTC CTT GAC CAA GGA GAG       432
78     Asp Ala Thr Ser Ile Leu Gly Ile Gly Thr Val Leu Asp Gln Gly Glu
79              130                 135                 140
80     ACT GCG GGG GCG AAA TTG GTT GTG TTC GCC ACC GCC ACC CCT CCG GGC       480
81     Thr Ala Gly Ala Lys Leu Val Val Phe Ala Thr Ala Thr Pro Pro Gly
82     145                 150                 155                 160
83     TCC GTC ACT GTG CCC CAT CCC AAC ATT GAG GAG GTT GCT CTA TCC ACC       528
84     Ser Val Thr Val Pro His Pro Asn Ile Glu Glu Val Ala Leu Ser Thr
85                  165                 170                 175
86     ACC GGA GAG ATC CCT TTT TAC GGC AAG GCT ATC CCC CTT GAG GTA ATC       576
87     Thr Gly Glu Ile Pro Phe Tyr Gly Lys Ala Ile Pro Leu Glu Val Ile
88                  180                 185                 190
89     AAG GGG GGG AGA CAT CTC ATC TTC TGT CAT TCA AAG AGG AAG TGC GAT       624
90     Lys Gly Gly Arg His Leu Ile Phe Cys His Ser Lys Arg Lys Cys Asp
91              195                 200                 205
92     GAG CTC GCC ACA AAG CTG GTC GCA ATG GGC ATC AAT GCC GTG GCC TAC       672
93     Glu Leu Ala Thr Lys Leu Val Ala Met Gly Ile Asn Ala Val Ala Tyr
94              210                 215                 220
95     TAC CGC GGT CTT GAC GTG TCC GTC ATC CCG ACC AGC GGT GAT GTT GTC       720
96     Tyr Arg Gly Leu Asp Val Ser Val Ile Pro Thr Ser Gly Asp Val Val
97     225                 230                 235                 240
98     GTC GTG GCA ACC GAC GCC CTC ATG ACC GGC TAT ACC GGC GAC TTC GAC       768
```

```
 99      Val Val Ala Thr Asp Ala Leu Met Thr Gly Tyr Thr Gly Asp Phe Asp
100                  245                 250                 255
101      TCG GTG ATA GAC TGC AAC ACG TGT GTC ACT CAG ACA GTC GAT TTC AGC        816
102      Ser Val Ile Asp Cys Asn Thr Cys Val Thr Gln Thr Val Asp Phe Ser
103                  260                 265                 270
104      CTT GAC CCT ACC TTC ACC ATT GAG ACG ACC ACA CTT CCC CAG GAT GCT        864
105      Leu Asp Pro Thr Phe Thr Ile Glu Thr Thr Thr Leu Pro Gln Asp Ala
106                  275                 280                 285
107      GTC TCC CGC ACT CAA CGA CGG                                            885
108      Val Ser Arg Thr Gln Arg Arg
109                  290                 295
111 (2) INFORMATION FOR SEQ ID NO: 2:
112        (i) SEQUENCE CHARACTERISTICS:
113            (A) LENGTH: 295 amino acids
114            (B) TYPE: amino acid
115            (D) TOPOLOGY: linear
116       (ii) MOLECULE TYPE: protein
117       (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:
118      Met Thr Met Ile Thr Asn Ser Arg Gly Ser Ile Met Lys Ser Pro Val
119       1               5                  10                  15
120      Phe Thr Asp Asn Ser Ser Pro Pro Val Val Pro Gln Ser Phe Gln Val
121                  20                  25                  30
122      Ala His Leu His Ala Pro Thr Gly Ser Gly Lys Ser Thr Lys Val Pro
123                  35                  40                  45
124      Ala Ala Tyr Ala Ala Gln Gly Tyr Lys Val Leu Val Leu Asn Pro Ser
125          50                  55                  60
126      Val Ala Ala Thr Leu Gly Phe Gly Ala Tyr Met Ser Lys Ala His Gly
127       65                  70                  75                  80
128      Ile Asp Pro Asn Ile Arg Thr Gly Val Arg Thr Ile Thr Thr Gly Ser
129                  85                  90                  95
130      Pro Ile Thr Tyr Ser Thr Tyr Gly Lys Phe Leu Ala Asp Gly Gly Cys
131                  100                 105                 110
132      Ala Gly Gly Ala Tyr Asp Ile Ile Ile Cys Asp Glu Cys His Ser Thr
133                  115                 120                 125
134      Asp Ala Thr Ser Ile Leu Gly Ile Gly Thr Val Leu Asp Gln Gly Glu
135          130                 135                 140
136      Thr Ala Gly Ala Lys Leu Val Val Phe Ala Thr Ala Thr Pro Pro Gly
137      145                 150                 155                 160
138      Ser Val Thr Val Pro His Pro Asn Ile Glu Glu Val Ala Leu Ser Thr
139                  165                 170                 175
140      Thr Gly Glu Ile Pro Phe Tyr Gly Lys Ala Ile Pro Leu Glu Val Ile
141                  180                 185                 190
142      Lys Gly Gly Arg His Leu Ile Phe Cys His Ser Lys Arg Lys Cys Asp
143                  195                 200                 205
144      Glu Leu Ala Thr Lys Leu Val Ala Met Gly Ile Asn Ala Val Ala Tyr
145          210                 215                 220
146      Tyr Arg Gly Leu Asp Val Ser Val Ile Pro Thr Ser Gly Asp Val Val
147      225                 230                 235                 240
148      Val Val Ala Thr Asp Ala Leu Met Thr Gly Tyr Thr Gly Asp Phe Asp
```

)

```
149                    245              250              255
150      Ser Val Ile Asp Cys Asn Thr Cys Val Thr Gln Thr Val Asp Phe Ser
151                260              265              270
152      Leu Asp Pro Thr Phe Thr Ile Glu Thr Thr Thr Leu Pro Gln Asp Ala
153            275              280              285
154      Val Ser Arg Thr Gln Arg Arg
155         290              295
157 (2) INFORMATION FOR SEQ ID NO: 3:
158        (i) SEQUENCE CHARACTERISTICS:
159              (A) LENGTH: 40 base pairs
160              (B) TYPE: nucleic acid
161              (C) STRANDEDNESS: single
162              (D) TOPOLOGY: linear
163       (ii) MOLECULE TYPE: cDNA
164       (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:
165        AAGGGATCCA TCATGAAATC CCCGGTGTTC ACGGATAACT                40
167 (2) INFORMATION FOR SEQ ID NO: 4:
168        (i) SEQUENCE CHARACTERISTICS:
169              (A) LENGTH: 39 base pairs
170              (B) TYPE: nucleic acid
171              (C) STRANDEDNESS: single
172              (D) TOPOLOGY: linear
173       (ii) MOLECULE TYPE: cDNA
174       (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:
175        GGGAAGCCTT AATTCTTACC GTCGTTGAGT GCGGGAGAC                 39
177 (2) INFORMATION FOR SEQ ID NO: 5:
178        (i) SEQUENCE CHARACTERISTICS:
179              (A) LENGTH: 39 base pairs
180              (B) TYPE: nucleic acid
181              (C) STRANDEDNESS: single
182              (D) TOPOLOGY: linear
183       (ii) MOLECULE TYPE: cDNA
184       (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:
185        GAGGGATCCA TCATGAAAGC GGTGGACTTT ATCCCTGTG                 39
187 (2) INFORMATION FOR SEQ ID NO: 6:
188        (i) SEQUENCE CHARACTERISTICS:
189              (A) LENGTH: 33 base pairs
190              (B) TYPE: nucleic acid
191              (C) STRANDEDNESS: single
192              (D) TOPOLOGY: linear
193       (ii) MOLECULE TYPE: cDNA
194       (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:
195        GAGAAGCTTT TAACACGTGT TGCAGTCTAT CAC                       33
197 (2) INFORMATION FOR SEQ ID NO: 7:
198        (i) SEQUENCE CHARACTERISTICS:
199              (A) LENGTH: 39 base pairs
200              (B) TYPE: nucleic acid
201              (C) STRANDEDNESS: single
202              (D) TOPOLOGY: linear
```

RAW SEQUENCE LISTING   DATE: 12/26/2001
PATENT APPLICATION: US/09/896,032   TIME: 15:29:49

Input Set : N:\Crf3\RULE60\09896032.txt
Output Set: N:\CRF3\12262001\I896032.raw

```
203     (ii) MOLECULE TYPE: cDNA
204     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:
205     GAGGGATCCA TCATGAAACA CCTGCATGCT CCCACCGGC                    39
207 (2) INFORMATION FOR SEQ ID NO: 8:
208     (i) SEQUENCE CHARACTERISTICS:
209         (A) LENGTH: 33 base pairs
210         (B) TYPE: nucleic acid
211         (C) STRANDEDNESS: single
212         (D) TOPOLOGY: linear
213     (ii) MOLECULE TYPE: cDNA
214     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 8:
215     GAGAAGCTTT TAATACCAAG CACAGCCTGC GTC                          33
217 (2) INFORMATION FOR SEQ ID NO: 9:
218     (i) SEQUENCE CHARACTERISTICS:
219         (A) LENGTH: 302 amino acids
220         (B) TYPE: amino acid
221         (D) TOPOLOGY: linear
222     (ii) MOLECULE TYPE: protein
223     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 9:
224     Pro Val Glu Asn Leu Glu Thr Thr Met Arg Ser Pro Val Phe Thr
225      1               5                  10                  15
226     Asp Asn Ser Ser Pro Pro Val Val Pro Gln Ser Phe Gln Val Ala
227                     20                  25                  30
228     His Leu His Ala Pro Thr Gly Ser Gly Lys Ser Thr Lys Val Pro
229                     35                  40                  45
230     Ala Ala Tyr Ala Ala Gln Gly Tyr Lys Val Leu Val Leu Asn Pro
231                     50                  55                  60
232     Ser Val Ala Ala Thr Leu Gly Phe Gly Ala Tyr Met Ser Lys Ala
233                     65                  70                  75
234     His Gly Ile Asp Pro Asn Ile Arg Thr Gly Val Arg Thr Ile Thr
235                     80                  85                  90
236     Thr Gly Ser Pro Ile Thr Tyr Ser Thr Tyr Gly Lys Phe Leu Ala
237                     95                  100                 105
238     Asp Gly Gly Cys Ala Gly Gly Ala Tyr Asp Ile Ile Ile Cys Asp
239                     110                 115                 120
240     Glu Cys His Ser Thr Asp Ala Thr Ser Ile Leu Gly Ile Gly Thr
241                     125                 130                 135
242     Val Leu Asp Gln Gly Glu Thr Ala Gly Ala Lys Leu Val Val Phe
243                     140                 145                 150
244     Ala Thr Ala Thr Pro Pro Gly Ser Val Thr Val Pro His Pro Asn
245                     155                 160                 165
246     Ile Glu Glu Val Ala Leu Ser Thr Thr Gly Glu Ile Pro Phe Tyr
247                     170                 175                 180
248     Gly Lys Ala Ile Pro Leu Glu Val Ile Lys Gly Gly Arg His Leu
249                     185                 190                 195
250     Ile Phe Cys His Ser Lys Arg Lys Cys Asp Glu Leu Ala Thr Lys
251                     200                 205                 210
252     Leu Val Ala Met Gly Ile Asn Ala Val Ala Tyr Tyr Arg Gly Leu
253                     215                 220                 225
```

VERIFICATION SUMMARY   DATE: 12/26/2001
PATENT APPLICATION: US/09/896,032   TIME: 15:29:51

Input Set : **N:\Crf3\RULE60\09896032.txt**
Output Set: **N:\CRF3\12262001\I896032.raw**

L:22 M:220 C: Keyword misspelled or invalid format, [(A) APPLICATION NUMBER:]
L:23 M:220 C: Keyword misspelled or invalid format, [(B) FILING DATE:]